Pre-lecture Problems for Lecture 7: Gradient Descent and Newton's Method

B10705034 資管三 許文鑫

2024 年 4 月 1 日

1. (10 points) Let's solve

$$\min_{x \in \mathbb{R}} f(x) = x_1^2 + x_2^2 - x_1 x_2$$

   (a) Find the gradient and Hessian of $f(x)$.

   Ans.

   gradient:

$$\begin{bmatrix} 2x_1 - x_2 \\ 2x_2 - x_1 \end{bmatrix}$$

   Hessian:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

   (b) Let $x^0 = (1, 2)$ be the initial solution. Run one iteration of gradient descent to get the next solution x^G. In that iteration, let the step size be that bringing you to the global minimum along the improving direction.

   Ans.

   - Step 0: $x^0 = (1, 2) . f(x^0) = 3$
   - Step 1:
     - $\nabla f(x^0) = (0, 3)$
     - $a_0 = \text{argmax}_{a \geq 0} f(x^0 - a\nabla f(x^0)), \text{where} f(x^0 - a\nabla f(x^0)) = (1, 2 - 3a) = 9a^2 - 9a + 3$ It follows that $a_0 = \frac{1}{2}$
     - $x^1 = x^0 - a_0 \nabla f(x^0) = (1, 2) - \frac{1}{2}(0, 3) = (1, \frac{1}{2})$. Note that $f(x^1) = \frac{3}{4}$
       $x^G = (1, \frac{1}{2})$

   (c) Let $x^0 = (1, 2)$ be the initial solution. Run one iteration of Newton's method to get the next solution x^F .

   Ans.

- let

$$f_Q(x) = f(x^0) + \nabla f(x^0)^T(x - x^0) + \frac{1}{2}(x - x^0)^T \nabla^2 f(x^0)(x - x^0)$$

  be the quadratic approximation of $f(x)$ at x^0. Note that we use the Hessian $\nabla^2 f(x^0)$

- We move from x^0 to x^1 by moving to the global minimum of the quadratic approximation:

$$\nabla f(x^0) + \nabla^2 f(x^0)(x^1 - x^0) = 0,$$

  i.e.

$$x^1 = x^0 - [\nabla^2 f(x^0)]^{-1} \nabla f(x^0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$x^F = (0, 0)$